Rehearsal and introduction

Equilibrium: $\pi = \pi \cdot P$

At this course we remain in the Markovian space. We’ll focus mostly on continuous Markov chains from now on.

$\lambda$ is the transition rate, and not the transition probability, and $p = \lambda \cdot \Delta t$.

We defined Q to be the transition rate matrix such that $\pi \cdot Q = 0$.

Birth-Death Chains:

The assumptions we need on a birth-death process are that in homogeneous Markov chain $X(t)$ where births and deaths are independent and

$B1: \Pr[\text{exactly 1 birth in } (t, t+\Delta t) \mid k \text{ in population}] = \lambda_k \cdot \Delta t + o(\Delta t)$

$D1: \Pr[\text{exactly 1 death in } (t, t+\Delta t) \mid k \text{ in population}] = \mu_k \cdot \Delta t + o(\Delta t)$

$B2: \Pr[\text{exactly 0 births in } (t, t+\Delta t) \mid k \text{ in population}] = 1 - \lambda_k \cdot \Delta t + o(\Delta t)$

$D2: \Pr[\text{exactly 0 deaths in } (t, t+\Delta t) \mid k \text{ in population}] = 1 - \mu_k \cdot \Delta t + o(\Delta t)$

While $o(\Delta t)$ is an expression describing a term that converges to 0 ‘faster’ than $\Delta t$. 
Look at the time Axis at $t + \Delta t$. If we want to be at $E_k$, what are all the possible states to be at at time $t$?

**Note:** when $t$ is infinitesimally small, only one death/birth can happen at $t+\Delta t$.

Denote $P_k(t + \Delta t)$ the probability to be at state $k$ at time $t + \Delta t$:

$$
P_k(t + \Delta t) = P_k(t) \cdot p_{kk}(\Delta t) + \left(P_{k-1}(t) \cdot p_{k-1,k}(\Delta t) + P_{k+1}(t) \cdot p_{k+1,k}(\Delta t) + o(\Delta t)\right)
$$

The above equation is for any $k > 0$. If we open the above terms we get:

$$
P_k(t + \Delta t) = P_k(t) \cdot [1 - \lambda_k \cdot \Delta t - \mu_k \cdot \Delta t + o(\Delta t)] +$$

$$P_{k-1}(t) \cdot [\lambda_{k-1} \cdot \Delta t + o(\Delta t)]$$

$$P_{k+1}(t) \cdot [\mu_k \cdot \Delta t + o(\Delta t)]$$

Changing sides and dividing by $\Delta t$ would result:

$$\frac{P_k(t + \Delta t) - P_k(t)}{\Delta t} = -P_k(t) \cdot \lambda_k - P_k(t) \cdot \mu_k + P_{k-1}(t) \cdot \lambda_{k-1} + P_{k+1}(t) \cdot \mu_{k+1} + o(\Delta t)$$

Adding $\lim$ to both sides results:

$$\frac{dP_k(t)}{dt} = -P_k(t) \cdot (\lambda_k + \mu_k) + P_{k-1}(t) \cdot \lambda_{k-1} + P_{k+1}(t) \cdot \mu_{k+1}$$

**Note:** when $k=0$, we need to zero the $k-1$ terms in the above equation.

The above describes the probability growth rate to be at state $k$. We see that the probability growth rate to be at any state is the sum of all ‘incoming’ rates minus the sum of all ‘outgoing’ rates.

We can ‘wrap’ any number of states and find out what the incoming and outgoing rates are, and retrieve the probability to be at those states.
**Birth Chains:**

Let’s examine a simpler process where $\forall k: \mu_k = 0$ - meaning birth-only process:

$$\frac{dP_k(t)}{dt} = -\lambda_k \cdot P_k(t) + \lambda_{k-1} \cdot P_{k-1}(t)$$

We can make the process even simpler by assuming $\forall k: \lambda_k = \lambda$ meaning the rates are constant. We now get:

$$\begin{cases}
\frac{dP_0(t)}{dt} = -\lambda \cdot P_0(t) \\
\frac{dP_k(t)}{dt} = -\lambda \cdot (P_k(t) - P_{k-1}(t)) \implies P_0(t) = e^{-\lambda t} \\
P_k(0) = \begin{cases} 1 & \text{if } k = 0 \\
0 & \text{otherwise}
\end{cases}
\end{cases}$$

where the above is a simple solution of differential equation. From that we can deduce $P_t(t)$:

$$\frac{dP_k(t)}{dt} = -\lambda P_k(t) + \lambda P_0(t) \implies P_k(t) = \lambda t \cdot e^{-\lambda t}$$

The global formula for $P_k(t)$ would then be:

$$P_k(t) = \frac{(\lambda t)^k}{k!} \cdot e^{-\lambda t}$$

This describes a **Poisson distribution function**; the probability of $k$ arrivals during interval $t$.

**Mean value $E[k]$**:

Consider the random variable $K$: the number of arrivals at interval $t$ (what used to be $\alpha(t)$). $K$’s mean value:

$$E[k] = \sum_{k=0}^{\infty} k \frac{(\lambda t)^k}{k!} e^{-\lambda t} = \lambda t \cdot \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t}$$

Remember the following holds: $\sum_{i=0}^{\infty} (\frac{\lambda t}{i!})^i e^{-\lambda t} = 1$. It’s easy to understand why if we think of each $i$ as the probability for exactly $i$ arrivals to happen at interval $t$. The summation of all probabilities of possible arrivals must give 1. Therefore, for every $t$:

$$E[k] = \lambda t$$

and the variance: $\sigma_k^2 = \lambda t$
**Moment generating function G(z):**

Let’s examine the $Z$-Transform of the Poisson probability function:

Define: $g_k = \Pr[K = k]$

$$G(Z) = \sum g_k \cdot Z^k \Rightarrow G(Z) = E[Z^k]$$

$$= \sum Z^k \cdot P_k(t) = \sum \frac{\left(\lambda t\right)^k \cdot Z^k}{k!} \cdot e^{-\lambda t} = e^{\lambda Z} \cdot e^{-\lambda t} \cdot \sum \frac{\left(\lambda t\right)^k \cdot Z^k}{k!} \cdot e^{-\lambda t} = e^{\lambda(Z - 1)}$$

Finally resulting $G(Z) = P_k$’s generating function:

$$G(Z) = E[Z^k] = e^{\lambda(Z - 1)}$$

By deriving $m$ times and assigning $z = 1$, we can calculate moment $m$ of $K$ specifically its mean and variance.

**Interarrival-times distribution**

We’d like to break $(0, t)$ to $2k+1$ time frames as shown in the figure below:

Consider the joint distribution of arrivals when it’s known beforehand that exactly $k$ arrivals have occurred during $(0, t)$.

**Define:** $A_k$ – [exactly 1 arrival in each of the $\beta_i$’s and no arrivals at in any of $\alpha_i$’s].

We’d like to calculate the probability of $A_k$ given exactly $k$ arrivals occurred in $(0, t)$:

$$\Pr[A_k | exactly \ k \ arrivals \ in \ (0, t)] = \frac{\Pr[A_k \ and \ exactly \ k \ arrivals \ in \ (0, t)]}{\Pr[exactly \ k \ arrivals \ in \ (0, t)]}$$

Note that the $\alpha$’s and $\beta$’s events are independent events, thus the probability can be calculated as a product of the individual probabilities:

$$\Pr[1 \ arrival \ in \ interval \ of \ length \ \beta_i] = \lambda \beta e^{-\lambda \beta_i}$$

And

$$\Pr[no \ arrival \ in \ interval \ of \ length \ \alpha_i] = e^{-\lambda \alpha_i}$$

Giving:

$$\Pr[A_k | exactly \ k \ arrivals \ in \ (0, t)] = \frac{(\lambda \beta_1 \cdot \lambda \beta_2 \cdots \lambda \beta_k \cdot e^{-\lambda \beta_1} \cdot e^{-\lambda \beta_2} \cdots e^{-\lambda \beta_k})(e^{-\lambda \alpha_1} \cdots e^{-\lambda \alpha_k})}{[(\lambda t)^k / k!]e^{-\lambda t}}$$

$$= \frac{\beta_1 \beta_2 \cdots \beta_k}{t^k} k!$$

*the last equation is since $t = \alpha_1 + \alpha_2 + \cdots + \alpha_k + \beta_1 + \beta_2 + \cdots + \beta_{k+1}$

Note that it doesn’t matter if the points on the time axis were picked randomly with uniform distribution, or picked in advance like in the above analysis.

Furthermore, it is easy to prove that the interval’s starting time doesn’t matter - only its length:

$$\Pr[X(s, s + \Delta t) = k] = \frac{k \cdot t^k}{k!} \cdot e^{-\lambda t}$$
\[ \tilde{t} \text{ exponential distribution:} \]

Define \( \tilde{t} \) - the time between adjacent arrivals.

Accordingly,

**Define the PDF and pdf** \( A(t) = \Pr[\tilde{t} < t] \) and \( a(t) = A'(t) \)

Calculation \( A(t) \) and \( a(t) \) is easy:

\[
A(t) = 1 - P[\tilde{t} > t] = 1 - P_{\tilde{t}}(t) = 1 - \lambda e^{-\lambda t}
\]

\[ a(t) = A'(t) = \lambda e^{-\lambda t} \]

This is the PDF and pdf of the exponential distribution – concluding:

Poisson process implies interarrival process with exponential distribution!

Specifically, this is a *memoryless* process whose properties were largely discussed at lecture-2:

\[
\Pr[\tilde{t} < t + t_0 \mid \tilde{t} > t_0] = \frac{\Pr[t_0 < \tilde{t} \leq t + t_0]}{P[\tilde{t} > t_0]} = \\
= \frac{\Pr[\tilde{t} \leq t + t_0] - \Pr[\tilde{t} \leq t_0]}{P[\tilde{t} > t_0]} = \\
= \frac{1 - e^{-\lambda (t+t_0)} - (1 - e^{-\lambda t_0})}{1 - (1 - e^{-\lambda t_0})} = \\
= 1 - e^{-\lambda t} = \\
= \Pr[\tilde{t} \leq t]
\]

\( \tilde{t} \) mean value:

\[
E[\tilde{t}] = \int_0^\infty a(t) \cdot t = \int_0^\infty \lambda e^{-\lambda t} \cdot t = -\lambda \frac{\partial}{\partial \lambda} \int_0^\infty e^{-\lambda t} dt = -\lambda \frac{\partial}{\partial \lambda} \left[ \frac{1}{\lambda} \right] = \lambda \cdot \frac{1}{\lambda^2}
\]

\( E[\tilde{t}] = \frac{1}{\lambda} \)
- **a constant birth rate**

Let’s finally examine the rate at an infinitesimal period $\Delta t$. Recall that:

$$\Pr[\tilde{t} < t_0 + \Delta t | \tilde{t} > t_0] = 1 - e^{-\lambda t_0} = 1 - [1 - \lambda \Delta t + \frac{(\lambda \Delta t)^2}{2!} + \ldots] = \lambda \Delta t + o(\Delta t)$$

Thus:

$$\Pr[\tilde{t} > t_0 + \Delta t | \tilde{t} > t_0] = 1 - \lambda \Delta t + o(\Delta t)$$

Moreover – the probability to have two or more arrivals in $(t_0, t_0 + \Delta t)$ is 0:

$$\Pr[\text{more than 1 arrival in } (t_0, t_0 + \Delta t)] = 1 - \Pr[0 \text{ arr in } (t_0, t_0 + \Delta t)] - \Pr[1 \text{ arr in } (t_0, t_0 + \Delta t)]$$

$$= 1 - (1 - \lambda \Delta t + o(\Delta t)) - [\lambda \Delta t + o(\Delta t)]$$

$$= o(\Delta t)$$

And so –

**Exponentially distribution implies constant birth rate $\lambda$.**

**Summary:**

We can summarize by concluding the relations between Poisson’s arrival process, the exponential interarrival times and the constant birth rate as depicted in the following figure. Each circle implies the other one!

For further usage we’ll calculate $A(t)$ Laplace transform:

$$A_S(a(t)) = \int_0^\infty e^{-a(t)} dt = \frac{\lambda}{\lambda + S}$$
**M/M/1 Birth-Death process:**
Consider a model where $\lambda_k = \lambda$ and $\mu_k = \mu$ - Poisson arrival process and Poisson service time:

\[ A(t) = 1 - e^{-\lambda t} \]
\[ B(t) = 1 - e^{-\mu t} \]

Note that the service time is the same as the death rate, since when finishing service the customer is gone.

From now and during chapter 3 we’re going to talk about the system after reaching equilibrium – the changes ‘rate’ is 0. Recall that $P_k(t)$ is the probability to have $k$ ‘customers’ at time $t$, the mathematical equivalent would be:

\[ \frac{dP_k(t)}{dt} = -(\lambda + \mu)P_k(t) + \lambda P_{k-1}(t) + \mu P_{k+1}(t) = 0 \]

This is the stationary state $\rightarrow t$ has no meaning since no changes occur in time anymore. For each state, the ‘enter-rate’ equals the ‘exit-rate’.

Consider the following diagram, focusing on the dashed cut between states $k$ and $k-1$.

The following holds:

\[ P_{k-1} \cdot \lambda = P_k \cdot \mu \]

and therefore:

\[ P_k = P_{k-1} \cdot \frac{\lambda}{\mu} \rightarrow P_k = P_0 \left( \frac{\lambda}{\mu} \right)^k. \]

Define \( \rho = \frac{\lambda}{\mu} \)

And we get that $P_k = P_0 \cdot \rho^k$.

As always, the following holds:

\[ \sum_{k=0}^{\infty} P_k = \sum_{k=0}^{\infty} P_0 \cdot \rho^k = 1 \]

Giving:

\[ P_0 = \frac{1}{\sum_{k=0}^{\infty} \rho^k} = 1 - \rho \]

Note that $P_0$ is the probability for the queue to be empty – meaning that $1-P_0$ is the probability for the queue to be full/active – which is why $\rho$ is the utility of the queue/server.

\[ P_k = \rho^k (1-\rho) \]
Chapter 3: Birth Death queueing Systems in Euqilibrium

What would change in case $\lambda_k \neq \lambda$ and $\mu_k \neq \mu$?
The equation should be changed to represent the dashed cut in the above figure:

$$P_k = P_{k-1} \frac{\lambda_{k-1}}{\mu_k} \Rightarrow$$

$$P_k = P_0 \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}$$

At some cases, this phrase converges to something we can calculate (for example when special relations between the different $\lambda_i$’s and $\mu_i$’s apply).

Convergence note: The above term converges iff starting some $k_i$, for every $i > k_0$: $\lambda_i > \mu_{i+1}$.
Note: the probabilities converge – the system does not converge to a certain state.

Calculating the queue size’s mean value $N$:

$$N = \sum_{k=0}^{\infty} k \cdot P_k = \sum_{k=0}^{\infty} k \cdot \rho^k (1 - \rho) = (1 - \rho) \rho \sum_{k=0}^{\infty} k \cdot \rho^{k-1} = (1 - \rho) \rho \cdot \frac{\partial}{\partial \rho} \sum_{k=0}^{\infty} \rho^k = (1 - \rho) \rho \cdot \frac{\partial}{\partial \rho} \left[ \frac{1}{1 - \rho} \right]$$

$$\bar{N} = \frac{\rho}{1 - \rho}$$

Recall Little Law: $N = \lambda T \Rightarrow T = \frac{\rho / \lambda}{1 - \rho} \Rightarrow T = \frac{1}{\mu} \frac{1 - \rho}{1 - \rho}$

Note that the average service time is $\frac{1}{\mu}$ and that $1 - \rho$ is 1-(utility) $\Rightarrow$ if $\rho = 0$ the customer will be in the system exactly the average service time as described in the following graph for M/M/1: